Planning Diagnostic Imaging Work-up Strategies using Case-Based Reasoning

Charles E. Kahn, Jr., M.D.

Section of Information and Decision Sciences, Department of Radiology,
Medical College of Wisconsin, Milwaukee, Wisconsin

ISIS is a developmental decision support system that helps physicians select diagnostic imaging procedures. It uses case-based reasoning, an artificial-intelligence approach that emphasizes reasoning and planning from prior experience. The development, training, and evaluation of a prototype system were used to guide the development of ISIS. To realize a clinically useful system, particular emphasis has been placed on increasing the depth and breadth of case-based knowledge, enhancing the explanatory capabilities of the system, and refining the human-computer interfaces to include a critiquing approach.

INTRODUCTION

How does a physician plan a patient's diagnostic imaging work-up? The physician might consider similar patients and the radiological procedures that were performed in those cases. Using this knowledge, the physician would decide which imaging procedure to request.

Case-based reasoning (CBR) is an approach to computer-based cognition that involves reasoning from prior experiences: new problems are solved by recalling and adapting solutions that were used to solve old problems [1,2]. CBR systems can be particularly useful in domains where well-defined causal models or extensive statistical data are lacking.

The knowledge base, or "memory," of a CBR system consists of cases indexed by their pertinent features. The dynamic operation of a CBR system involves: (1) encoding new cases and storing them into memory, (2) activating (retrieving) cases from memory that are pertinent to the current situation, and (3) adapting the actions of retrieved cases to compute a course of action for the current situation.

ISIS (Intelligent Selection of Imaging Studies) is a case-based decision support tool being developed to help physicians select appropriate radiological procedures [3]. It provides computer-based expertise in the domain of diagnostic imaging procedures such as computed tomography (CT), ultrasound (US), and magnetic resonance imaging (MRI). ProtoISIS, a prototype case-based system, learned the use of ultrasonography and body CT from 200 actual, consecutive requests from imaging procedures, from which it achieved 84% accuracy in classifying new cases [3]. This report discusses the directions taken in the development of ISIS based on lessons learned from the clinical trial and evaluation of ProtoISIS [4].

PROTOTYPE SYSTEM

Exemplar-Based Learning

ProtoISIS is based on a CBR shell called Protos. Protos performs case-based classification: it learns to classify cases based on associations between categories and exemplary cases ("exemplars") [5,6]. Protos attempts to classify a new case by matching it to cases with similar features. We implemented CL-Protos, a version of Protos in the Common Lisp language [7], in Macintosh Common Lisp 2.0 on Macintosh IIsi and PowerBook 180 computers (Apple Computers, Cupertino, Calif.).

An "exemplar" is a case that particularly represents the specified category. Each exemplar in Protos consists of a name, a set of features and a classification. Each term known to the system may have an abbreviation and one or more synonyms. This information is supplemented by explanations that relate two or more terms. When a new case is presented, Protos gives the user the choice to pre-classify the case or to let Protos suggest a classification. If no suitable case is found, Protos asks the user to classify the case and to provide an explanation that relates the features of the case to the specified category.

Terms and Relations

In ProtoISIS, a case consists of a request for an imaging procedure. The order number is the case

name, the clinical indications and questions to be answered are the features of the case, and the imaging study performed is the classification. Each semantically descriptive item — case name, feature, or category — is a "term"; terms can have synonyms, abbreviations, and relations to other terms.

To establish conceptual relationships between terms such as features and categories, ProtoISIS elicits explanations from the user. These relationships allow ProtoISIS to distinguish new cases from previously learned exemplars, and to link the case's features to its classification. Subsumption ("is a"), causal, functional, definitional, and part/whole relations can be defined.

For example, case PT070 has features FATTY-FOOD-INTOLERANCE and RIGHT-UPPER-QUADRANT-PAIN. The imaging-procedure category for this case is GALLBLADDER-ULTRASOUND; this term has the synonym US-GALLBLADDER and the abbreviation US-GB. In addition, ProtoISIS knows that GALLBLADDER-ULTRASOUND is a specialization of ULTRASOUND-PROCEDURE, which is in turn a specialization of IMAGING-PROCEDURE.

Because the original Protos system did not contain a verb to express the concept "detects" or "reveals," the verb VISUALIZES was substituted for ENABLES to relate imaging procedures and the conditions they reveal. Relations can be qualified by the terms ALWAYS, USUALLY, SOMETIMES, and OCCASIONALLY. For example, to the question "How is RIGHT-UPPER-QUADRANT-PAIN related to US-ABDOMEN?", one could answer: "US-ABDOMEN USUALLY VISUALIZES GALLSTONES" and "GALLSTONES SOMETIMES CAUSES RIGHT-UPPER-QUADRANT-PAIN."

These explanations allow ProtoISIS to create a semantic network that relates the cases, features, and imaging procedures (Figure 1). In addition, ProtoISIS creates "remindings" from the case's features to its imaging-procedure category, such as from RIGHT-UPPER-QUADRANT-PAIN to US-ABDOMEN, if it judges the feature's association with the category to be sufficiently strong. Strength is determined in
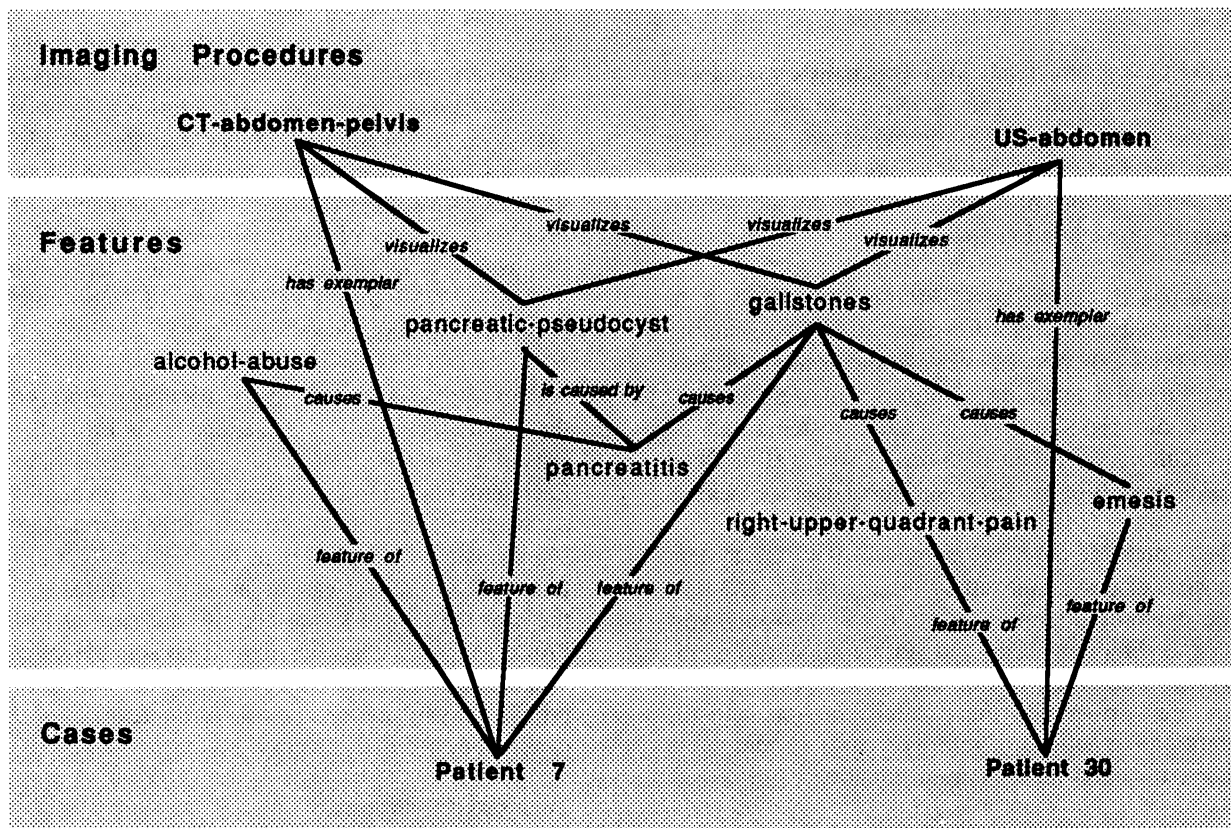

Figure 1. Semantic network of terms created by ProtoISIS.

part by the verbs (e.g., CAUSES, DETECTS or DEFINITION IMPLIES) and the qualifiers (e.g., ALWAYS, USUALLY) that express the relationships that lead from the feature to the category. These remindings allow ProtoISIS to associate features of new cases with relevant categories.

Performance of Prototype System

ProtoISIS was trained with 200 actual cases of body CT and ultrasound requests from one week of radiology department records. To test the system, 100 new, consecutive ultrasound and body CT cases, grouped into four sets of 25 cases each, were presented sequentially. After each case's identifier and clinical features were entered, the system attempted to assign the correct category to each case. If ProtoISIS was unable to assign a category or assigned an incorrect category to a case, we added that case and pertinent explanations into memory. ProtoISIS incorporated into its knowledge base all new terms — such as abbreviations, synonyms, and features — that were encountered in the test cases whether or not the case to which they belonged was itself added.

After training, ProtoISIS incorporated a total vocabulary of 527 terms: 200 case names, 28 imaging procedures, 37 abbreviations, 40 synonyms, and 222 features. Of the nine CT procedures, CT-ABDOMEN-PELVIS, CT-CHEST-ABDOMEN-PELVIS, and CT-CHEST had the most exemplars (33, 22, and 16, respectively). Among the 19 ultrasound procedures, US-KIDNEY, US-ABDOMEN, DOPPLER-ABDOMEN and

US-HEAD had the most exemplars (29, 26, 14, and 11, respectively). All other imaging-procedure categories had six or fewer exemplars. Each exemplar consisted of one or more features: 66 exemplars (33%) had only one feature, another 77 (38%) had two features, and none had more than seven features. All but 13 (5.9%) of the 222 features had remindings to one or more imaging procedures. The great majority of features (76.4%) had remindings to only one imaging procedure; none had remindings to more than three imaging procedures.

In classifying imaging-procedure requests, ProtoISIS demonstrated satisfactory performance (Figure 2). Only three of the 100 test cases required new imaging-procedure categories: CT-ABDOMEN-DRAINAGE, CT-LIVER-BIOPSY, and US-AORTA. Overall, ProtoISIS correctly classified 72% of the imaging-procedure requests on the first attempt. Its performance improved as it gained experience: in the last two test series, it correctly classified 84% of the cases presented, compared with only 56% in the first series.

In many of the incorrectly classified cases, the correct imaging procedure received the second highest matching score. On average, 40% of cases included terms that had not been encountered previously; given the small number of training cases, the large vocabulary of medicine, and the variety of ways that a single medical concept can be expressed, this finding is not surprising.
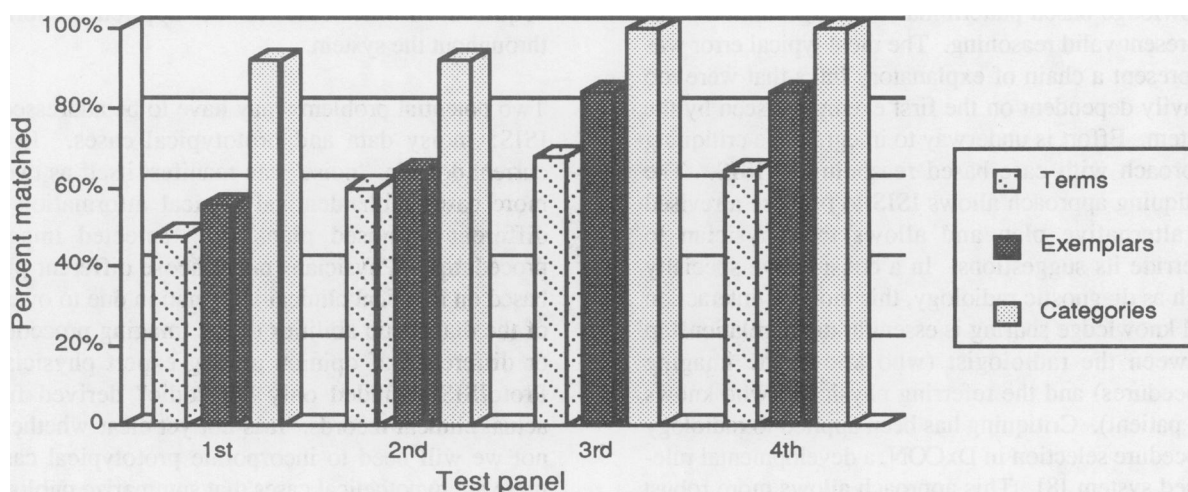

Figure 2. Performance of ProtoISIS. "Terms": cases that did not require additional terms or explanations. "Exemplars": cases that successfully matched an existing exemplar and were categorized correctly. "Categories": cases that did not require an additional category.

933

PRODUCTION SYSTEM

Case Structure and Reasoning
ISIS builds on the framework of ProtoISIS and incorporates two additional properties to overcome the prototype's deficiencies. First, ISIS distinguishes between known features (patient history) and those being queried (clinical questions). In addition to the procedure requested and the clinical information provided, each case includes information about the procedure actually performed, the imaging technique or protocol, and clinical questions to be asked of the referring physician. Second, ISIS treats imaging procedures as a elements of a plan, rather than as categories. ISIS can modify a plan's components instead of creating a unique category for each imaging protocol. This approach facilitates proper sequencing of imaging procedures and offers much richer interaction between computer and physician.

Interaction with Physicians
Although CL-Protos provides efficient and easily understandable interfaces, evaluation of their use in ProtoISIS found them unsuitable for physicians. In ISIS, the user interfaces limit the functionality of the system to those features needed by physicians. In addition, information being entered or presented must be clustered in ways that corresponds to typical clinical scenaria. Because most commercial radiology information systems are based on VT100-type 24-line, 80-column displays, integration with such systems requires special attention to the user interfaces to assure ease of use.

Some of the explanations generated by the Protos knowledge-based pattern matching algorithm did not represent valid reasoning. The most typical error was to present a chain of explanatory links that were too heavily dependent on the first exemplars seen by the system. Effort is underway to integrate the critiquing approach with case-based reasoning in ISIS. The critiquing approach allows ISIS to propose a revised or alternative plan and allows the physician to override its suggestions. In a consultative specialty such as diagnostic radiology, this mode of interaction and knowledge sharing is essential to the relationship between the radiologist (who knows the imaging procedures) and the referring physician (who knows the patient). Critiquing has been applied to radiology procedure selection in DxCON, a developmental rule-based system [8]. This approach allows more robust interaction between the physician and the decision support system.

DISCUSSION

As medical imaging technology has advanced, selecting appropriate diagnostic imaging procedures has become increasingly complex. Collaborative planning of imaging work-up strategies by radiologists and referring physicians has been limited by the large number of consultation requests. A computer-based "proxy" for the radiologist could help referring physicians select optimal diagnostic imaging procedures.

Existing decision support systems for radiological procedure selection include rule-based systems [8-10], hypertexts [11,12], and belief networks [13]. The development of ProtoISIS demonstrated that case-based reasoning can be applied successfully to the task of selecting diagnostic imaging procedures and that such a CBR system can be much simpler to construct and validate.

CBR systems offer other advantages. Knowledge acquisition (learning) for CBR systems is uncomplicated. Existing cases can be used to train the system; a causal model or deep understanding of the domain is not required. Also, human experts often have greater ease distilling their knowledge into examples rather than rules. CBR systems can explain their reasoning by referring to prior cases and generalizations of these cases. They can operate under incomplete knowledge and can adapt with experience; this flexibility is especially important in rapidly evolving domains such as medicine. CBR systems are computationally efficient and their network of relations between terms allows knowledge acquired in one case to be applied generally throughout the system.

Two potential problems may have to be addressed in ISIS: noisy data and prototypical cases. In the current domain, "noise" can manifest itself as two or more cases with identical clinical information but different proposed plans (*i.e.*, selected imaging procedures). Physicians may choose different plans based on identical clinical information due to overlap of the diagnostic abilities of the imaging procedures or difference of opinion among expert physicians. ProtoISIS included only "episodic" derived from actual clinical records. It is not yet clear whether or not we will need to incorporate prototypical cases; these are hypothetical cases that summarize published data or medical practice guidelines.

Case-based reasoning has been applied experimentally in medicine to diagnosis of hearing disorders [5,6,14], diagnosis of heart failure [15,16], and planning of radiation therapy protocols [17,18]. Although these systems have undergone some degree of validation, none is in routine clinical use. Once ISIS has been completely validated, it will integrated with our department's radiology information system, where it will provide interactive, on-line expertise to physicians at all times of the day, and be available to physicians in their work areas, such as clinics, inpatient wards, intensive-care units and the emergency department. Our radiology department performs more than 200,000 procedures annually, including about 36,000 imaging procedures. This setting will provide an excellent "production system" test, and will serve as a pilot project for "scaling up" the system to include all radiology procedures. ISIS offers an opportunity to study the role of case-based reasoning in day-to-day medical decision making, and has potential to significantly improve the quality and cost-effectiveness of medical care.

References
[1] Riesbeck C, Schank R. Inside Case-Based Reasoning. Hillsdale, NJ: Lawrence Erlbaum, 1989.

[2] Kolodner JL. An introduction to case-based reasoning. Artif Intel Rev 1992; 6:3-34.

[3] Kahn CE Jr, Anderson GM. Case-based reasoning and imaging procedure selection. Invest Radiol 1994; 29:643-647.

[4] Kahn CE Jr. Clinical trial and evaluation of a prototype case-based system for planning medical imaging work-up strategies. In: Aha DW, ed. Case-Based Reasoning: Papers from the 1994 Workshop. Menlo Park, CA: AAAI Press, 1994: 138-142.

[5] Bareiss ER, Porter BW, Wier CC. PROTOS: an exemplar-based learning apprentice. Int J Man-Machine Studies 1989; 29:549-561.

[6] Porter BW, Bareiss R, Holte RC. Concept learning and heuristic classification in weak-theory domains. Artif Intel 1990; 45:229-263.

[7] Dvorak D. CL-Protos User's Guide. Austin, TX: Artificial Intelligence Laboratory, Department of Computer Sciences, The University of Texas at Austin, 1989.

[8] Swett HA, Rothschild M, Weltin GG, Fisher PR, Miller PL. Optimizing radiologic workup: an artificial intelligence approach. J Digit Imaging 1989; 2:15-20.

[9] Kahn CE Jr, Messersmith RN, Jokich MD. PHOENIX: an expert system for selecting diagnostic imaging procedures. Invest Radiol 1987; 22:978-980.

[10] Kahn CE Jr. Validation, clinical trial, and evaluation of a radiology expert system. Methods Inf Med 1991; 30:268-274.

[11] Greenes RA, Tarabar DB, Krauss M, et al. Knowledge management as a decision support method: a diagnostic workup strategy application. Comput Biomed Res 1989; 22:113-135.

[12] Kahn CE Jr. A radiology hypertext system for education and clinical decision making. J Digit Imaging 1991; 4:207-212.

[13] Haddawy P, Kahn CE Jr, Butarbutar M. A Bayesian network model for radiological diagnosis and procedure selection: work-up of suspected gallbladder disease. Med Phys 1994; 21:1185-1192.

[14] Bareiss R. The experimental evaluation of a case-based learning apprentice. In: Hammond K, ed. Proceedings of a Workshop on Case-Based Reasoning. San Mateo, CA: Morgan Kaufmann, 1989: 162-167.

[15] Koton P. A medical reasoning program that improves with experience. Comput Methods Programs Biomed 1989; 30:177-184.

[16] Koton P. Evaluating case-based problem solving. In: Hammond K, ed. Proceedings of a Workshop on Case-Based Reasoning. San Mateo, CA: Morgan Kaufmann, 1989: 173-175.

[17] Berger J. ROENTGEN: case-based reasoning and radiation therapy planning. In: Proceedings of the 16th Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1992: 210-214.

[18] Berger J. Knowledge acquisition and design support in a medical domain. In: Leake DB, ed. Case-Based Reasoning: Papers from the 1993 Workshop. Menlo Park, CA: AAAI Press, 1993: 141-146.